

# Calculating Trade Damages in the Context of the World Trade Organization's Dispute Settlement Process

Jason Bernstein and David Skully

---

Since its inception in 1995, the World Trade Organization (WTO) dispute settlement system has received over 250 notification of trade disputes. While most have been settled, in a few cases the WTO arbitrators had to approve damage awards. This paper will use one of these cases (*Hormones*) to explain the methodology arbitrators use to calculate damages and how this methodology differs from an efficiency-base measure of welfare that economists would instinctively provide. Yet, there are rational reasons for this difference and the arbitrator's methodology does a better job of providing incentives for countries not to violate trade agreements.

---

One of the major achievements of Uruguay Round was the establishment of a new dispute settlement system for the World Trade Organization (WTO). Since its inception in 1995, the system has received over 250 notification of trade disputes involving distinct matters. However, only a few cases have gone through the final stage—retaliation—where the Dispute Settlement Body (DSB) awards damages for a country's failure to comply with its obligations under the WTO agreements. Two of these cases involved agricultural trade and received much media attention in the late 1990s—the *Bananas* dispute and the *Hormones* dispute, both involving the United States and the European Union (EU).

The DSB's job is arbitration. We argue that arbitration requires a different kind of economic analysis than what we, as international trade economists, would instinctively provide. An arbitrator's measure of trade impairment (damage) is not commensurate with an economist's efficiency-base measure of deadweight welfare loss. The DSB estimates trade impairment as foregone sales; it is a rectangle of foregone revenue, not a triangle of foregone welfare. There are, however, very rational, nonpolitical reasons for this difference. The

■ Jason Bernstein and David Skully are economists at the Economic Research Service of the United States Department of Agriculture.

*Hormones* case illustrates the kind of economic arguments and evidence effective in arbitration.

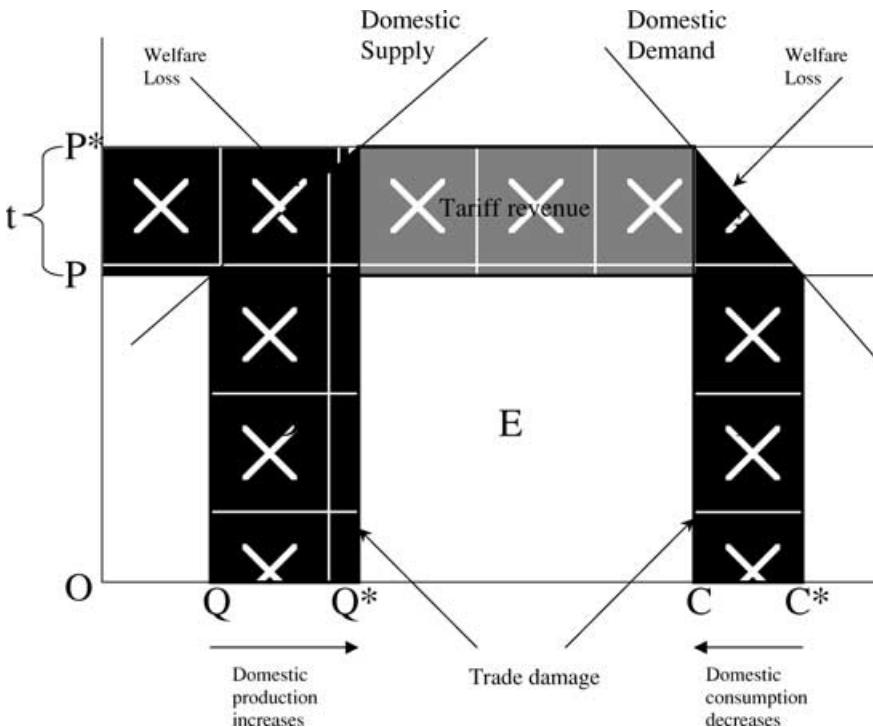
**Economists Think Triangles; Arbitrators Think Rectangles**

Arbitrators and economists use different measures of damages because they pose different questions: Economists are concerned about efficiency, while arbitrators are concerned about dispute resolution and distributional justice. Arbitrators and economists do not differ as to the facts of the case: They differ as to which facts to emphasize.

In trade disputes, arbitrators and economists wish to determine the magnitude of the alleged damage or trade impairment. The analytical problem is to compare the world with the policy to its counterfactual, the world without the policy. Consider the imposition of a specific tariff,  $t$ , on a freely traded good (figure 1).<sup>1</sup> The tariff does not affect the world price of the good,  $P$ , but it increases the domestic price to  $P^* = P + t$ . The higher domestic price encourages domestic firms to increase production from  $Q$  to  $Q^*$  and causes domestic consumers to reduce consumption from  $C$  to  $C^*$ . The volume of duty-free imports is  $QC$  and its value is the area  $D + E + F$ . Import volume with the tariff is  $Q^*C$  and its value is the area  $E$ . The tariff reduces import value by  $D + F$ .

The economic analysis of the welfare or efficiency loss from the tariff identifies income transfers between sectors (rectangles) and dead-weight welfare losses

**Figure 1. Partial-equilibrium analysis of a tariff (small country)**



(triangles). The tariff reduces the welfare of domestic consumers: Their loss of surplus is the area  $A + B + \text{tariff revenue} + G$ .  $A + B$  is a transfer from consumers to producers, “tariff revenue” is a transfer from consumers to the government, and rectangle  $G$  is a net welfare loss. The protected industry benefits from the tariff: Revenue increases by the area  $A + B + D$ .  $B + D$  is payment to factors of production, leaving a net surplus, or rent, equal to  $A$ . The tariff contrives an opportunity for the less efficient domestic industry to supply output  $QQ^*$  at a higher resource cost than foreign firms. Triangle  $B$  is thus a net welfare loss.

The net welfare or efficiency loss caused by the tariff is the sum of triangles  $B$  and  $G$ . All other effects are transfers between domestic agents and, from a national welfare perspective, they net out. The answer  $B + G$  will not resolve a trade dispute, however. The dispute concerns the unrequited and unilateral transfer of resources from one trading partner to another. The damage is the reduction in trade caused by the tariff. The arbitrator’s measure of damage is the sum of rectangles  $D$  and  $F$ , an area far larger than  $B + G$ . From a law-enforcement perspective, the economist’s measure,  $B + G$ , is not sufficient to deter a revenue-seeking and rent-creating government from imposing the tariff. Nor would a fine of  $B + G$  likely be sufficient to negate the realized gains of  $A + \text{tariff revenue}$ .

If the economic measure of welfare loss were employed, there would be little downside consequence to violating WTO obligations.<sup>2</sup> Nor would there be much incentive for injured parties to initiate disputes within the WTO’s dispute settlement process. The credibility of the WTO and its dispute settlement mechanism would collapse. Judicial bodies have devised means of getting around this problem as the next section illustrates.

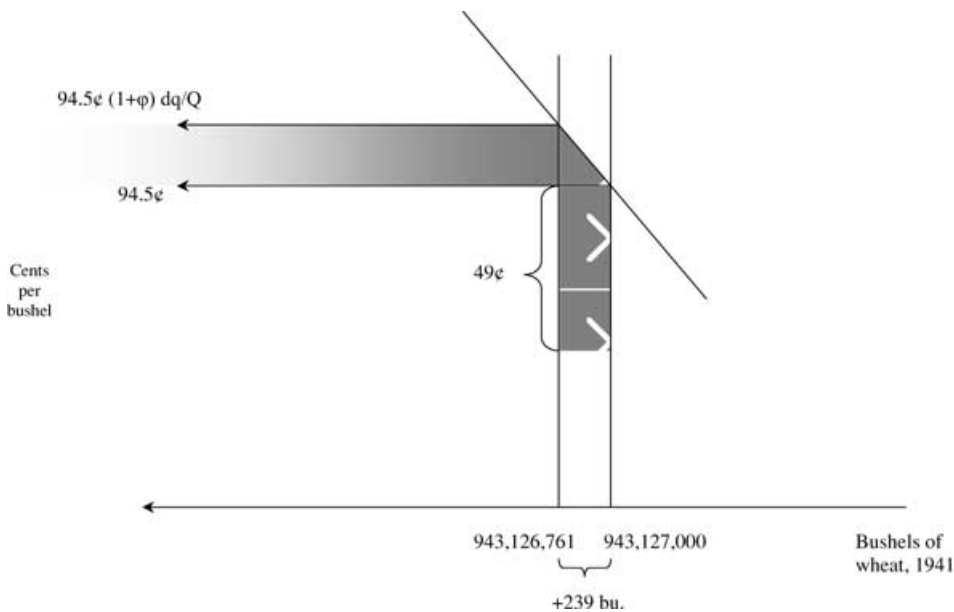
### The Crime of Roscoe C. Filburn

In 1941, Roscoe C. Filburn ignored USDA wheat allotment regulations and planted 11.9 more acres than allowed. From these acres, Filburn harvested 239 bushels of wheat and was fined 49 cents per bushel, a sum of \$117.11. Filburn refused to pay. After numerous appeals, the case reached the U.S. Supreme Court, which ruled in the USDA’s favor.

Filburn’s wheat did not leave his farm: He used it for seed and feed. He asserted that what he produced and consumed on his private property did not concern the federal government. The USDA countered that changes in on-farm consumption of wheat was the major source of variability in the marketed supply of wheat. What was true in aggregate was also true of the individual farm. The Court concurred:

That [Filburn’s] own contribution to the demand for wheat may be trivial by itself is not enough to remove him from the scope of federal regulation where, as here, his contribution, taken together with that of many others similarly situated, is far from trivial.<sup>3</sup>

This is the doctrine of aggregate or cumulative effect. It poses the hypothetical—what if everyone did that? If other farmers did as Filburn had done, it would have undermined the wheat production restriction program; therefore no farm should be exempt simply because the wheat remained on-farm.<sup>4</sup> The Court determined that Filburn’s use of the marginal bushel of wheat on his farm was an act of interstate commerce and thus subject to federal regulation. *Wickard v. Filburn*

**Figure 2. Triangles and rectangles in *Wickard v. Filburn***

remains on Constitutional Law reading lists because the decision allowed a very broad interpretation of the scope of the interstate commerce regulatory powers of the federal government.<sup>5</sup>

Filburn's crime was to produce 239 excess bushels of wheat, making the 1941 harvest 943,126,761 bushels rather than 943,127,000. Filburn increased the supply of wheat harvested by 0.000025%, a trivial change by any standard.

The "damage" caused by Filburn's action can be approximated by the shaded triangle in figure 2. The additional 239 bushels infinitesimally increased supply and reduced price. For any range of elasticities, consumers gain more than producers lose so welfare increases though the domestic wheat allotment program had nothing to do with welfare maximization or allocative efficiency; the allotments were imposed to increase commodity prices and farm income.

There are two hypotheses about the economic damage caused by Filburn's actions and the penalty imposed: Each hypothesis corresponds to a shaded rectangle in figure 2. The *horizontal hypothesis* is that Filburn's excess production imposed economic damages on all other wheat producers by increasing the supply of wheat by 0.000025% and reducing the price received by other producers proportionately. The damage equals the area of the shaded "horizontal" rectangle. Our upper bound estimate of the price change is only 0.00010%; but when multiplied by 943,126,761 bushels, the result is not trivial.<sup>6</sup>

The *vertical hypothesis* is that the fine is a *penalty* to deter free riding on the wheat allotment program. The 49-cent-per-bushel fine was half of the 1941 parity loan rate for wheat. The fine was not calculated to be an estimate of the horizontal rectangle of damages; rather it was a sufficiently large unit penalty to induce program compliance. The penalty is roughly half the area of the "vertical" rectangle.<sup>7</sup>

The vertical or penalty hypothesis is more plausible than the horizontal or damage hypothesis for the Filburn case. The unit fine was established *ex ante* to induce program compliance, that is, it was a penalty. Filburn's fine was not based on an *ex post* estimate of trade damages, nor was the fine revenue used as a basis for compensation to the damaged parties, as in a WTO dispute or a civil case.

The Filburn case shows that for the administration of justice to be credible, the compensation-granted injured parties must be sufficient to deter potential offenders—it must serve as a prospective penalty. Compensation must also be sufficient to induce injured parties to pursue compensation claims, but not so great as to induce opportunistic, nuisance cases.

Of the three areas—the horizontal rectangle, the triangle, and the vertical rectangle—the vertical rectangle is the easiest to determine *ex post*. It is the product of the market price, which can be observed, and the volume of trade impairment, which can be estimated by the construction of the appropriate counterfactual. The other two areas require an elasticity. One way to inhibit settlement and prolong a dispute is to introduce arguments about the “appropriate elasticity.” Arguments about elasticities are digressions: Arbitrators avoid them. What arbitrators want is a straightforward method of calculating damages based on the difference between the actual market situation and the situation that would have occurred if the offending measure were withdrawn.

## Weighing Evidence

The example above shows that arbitrators view damages differently than economists. This is not to say that all arbitrators or all economists think alike. However, arbitrators are more likely to view damages as a penalty measure to discourage further violations of the law while economists regard damages as a measure of the net welfare loss caused by the violation in question. Since a damage award equal to a measure of lost welfare would not be rational from a legal perspective, an alternative is to estimate the amount of economic damages caused by the violation in question. *Wickard v. Filburn* showed that there is a way to calculate a penalty that can discourage future violations of the law and, at the same time, approximate the economic damages caused by the violation.

This still leaves the question as to how arbitrators actually determine this penalty from the damage estimates provided by the parties to the dispute. The DSB panel operates under the rules of conventional arbitration (in contrast to “final-offer arbitration”) where the panel simply renders a decision that represents their assessment of the level of impairment. The settlement may be a compromise between the parties' final offers. Arbitrators cannot award punitive damages and are under no obligation to base their calculations on decisions in previous cases. DSB panels request both written and oral presentations of the disputants' estimates, including a description of the methodology used. The DSB, in turn, asks questions of their own (both during and after the hearing) which disputants are expected to answer.

One theory of conventional arbitrator behavior is that arbitrators determine the final damage award by a simple mechanical compromise between the parties' final offers (Benson; Cooter and Rubinfeld). Bloom finds that conventional

arbitrators tend to mechanically compromise between the parties' final offers, with virtually no evidence of systematic reference to the facts of the case. This strategy of "splitting the difference" between parties' estimates is not naïve: It may be optimal for arbitrators who want to appear impartial. The WTO and other multilateral bodies may be particularly sensitive to such criticisms. A mechanical compromise is also easier and less time-consuming than weighing the facts in each case, particularly in complex trade cases such as *Bananas* and *Hormones*.

However, when final offers convey information about the facts of the case, "splitting the difference" ceases to be a plausible description of arbitrator behavior.<sup>8</sup> Farber suggests that arbitrators use final offers as a source of information about facts. In a study of awards by 64 professional arbitrators in 25 simulated cases, Bazerman and Farber found that arbitrators weigh facts more heavily than offers. The facts become increasingly important as the offers diverge, contradicting "the naïve split-the-difference view of arbitrator behavior" (p. 76).

Of course, this still does not explain how arbitrators in general and the DSB in particular weigh the facts and offers to determine the level of damages. One hypothesis is that arbitrators determine the final damage as a weighted average of the parties' offers. Each disputant faces a trade-off between the magnitude and credibility of its offer. Plaintiffs have an incentive to inflate claims, but not so much as to undermine credibility. Similarly, defendants have an incentive to low-ball estimates, but they cannot be so low as to lack credibility. Arbitrators hear the two presentations and query both sides on points of fact and logic. The panel then adjourns to weigh the evidence and determine a just settlement. Each claim is decomposed into subclaims, each the product of an assertion and a monetary value. Some assertions are clearly valid or invalid; others have plausibility. The panel discards the invalid claims and weighs the value and plausibility of the remaining claims to reach a defensible settlement value.

Under this hypothesis, arbitrators consider the facts of the case as presented to them, but weigh the plausibility of the facts against the parties' offers. The hypothesis also offers a rationale of how it could *appear* that the arbitrators are simply "splitting the difference" to calculate damages. If both parties have an equal tendency to over- and underestimate their offers respectively (i.e., the weights are about equal), then a weighted average would approximate the simple average of the parties' offers. That arbitrators' damage awards approximate the average of the disputants' claims is an incentive for the parties to settle rather than pursue arbitration.

Disputants usually decide to continue to litigate—not settle—if their respective estimates of the damage are far apart, that is, if the expected gain from further litigation exceeds its cost (Posner). If disputants face similar litigation costs, a good way to influence a settlement would be to make the plaintiff's and defendant's estimate of the damage equal, or at least to reduce the difference between the two estimates. Therefore, arbitrator decisions may appear to be mechanical compromises of the parties' final positions, but only because the parties aligned themselves around the arbitrator's preferred settlement point (Bloom; Ashenfelter and Bloom; Farber; Gunderson; Ashenfelter). While arbitrators may not explicitly calculate the final damage award as the average of the two parties' estimates, such a methodology will result in most cases being settled.



The next section explains how the DSB arbitrators calculated damages in the *Hormones* case (hereafter referred to simply as *Hormones*.) *Hormones* is unique in that it describes the methodology the DSB panel used to determine the final damage number and provides a summary of how the panel weighted the disputants' claims.

### **The *Hormones* Dispute**

In 1989, the European Commission (EC) imposed an EU-wide ban on the imports of all animals and meat treated with growth-promoting hormones. This restricted almost all imports of High-Quality Beef (HQB) and Edible Beef Offal (EBO) from the United States and Canada. The EC contended that it enacted the ban to protect consumer health and safety. The United States and Canada claimed the ban violated the Sanitary and Phytosanitary (SPS) Agreement: It was not based on scientific principles, on risk assessments, or on guidelines for acceptable risks. Under the pre-Uruguay Round dispute settlement system (Article XXIII), the EU was able to block dispute settlement. The United States retaliated with restrictions against imports of EU products (Stewart).

In 1996, the United States and Canada requested WTO-based consultations with the EU, the first stage of the then newly formed WTO dispute settlement process. After attempts at settlement failed, a DSB panel report in August 1997 found that the EU hormone ban did not comply with the SPS Agreement and the EU's WTO obligations. A DSB appellate panel report in January 1998 upheld this finding and gave the EU 15 months to comply. In May 1999, after the 15-month period, the EU announced it would not remove the ban. In June 1999, the United States and Canada asked the WTO for authorization to suspend tariff concessions to the EU in the amounts of US\$202 million and C\$75 million, respectively. The EU countered that impairment was US\$53.3 million for the United States and C\$3.5 million for Canada. In July, the WTO authorized the United States and Canada to suspend concessions of US\$116.8 and C\$11.3 million, respectively.

### **General Principles Underlying the DSB's Calculations**

#### ***Gross Trade***

The DSB estimated the *gross* value of U.S. beef exports to the EU (both HQB and EBO) impaired by the ban. This is similar to the vertical rectangle in figure 2. The rectangle overestimates the damage to *total* U.S. export revenue, which is similar to the triangle in Figure 2. The EU consumed a small portion of the world HQB supply because of a preexisting quota on HQB imports. However, the EU consumed a sufficiently large share of world EBO trade to affect world prices, so the ban reduced total EBO demand and reduced the world price. This encouraged EBO imports in other markets, particularly in Asia. Although much of the export revenue lost in the EU market was regained from sales to third countries, the DSB did not consider this fact in the *Hormones* case. As in *Wickard v. Filburn*, the rectangle methodology was used in the *Hormones* case because it is simple, straightforward, and, while not sophisticated, does a reasonable job of approximating the amount of impairment caused by the ban.

## **Bilateral Trade**

The DSB only considers bilateral trade damages imposed on the complainant by the defendant: No third country effects are considered. The DSB emphasizes this point in *Bananas* (European Communities, 1999b)

We are of the view that the benchmark for the calculation of nullification or impairment of US trade flows should be losses in US exports of goods to the European Communities. . . . However, we are of the opinion that losses of US exports in goods or services *between the US and third countries* do not constitute nullification or impairment of even *indirect* benefits accruing to the United States under the GATT or the GATS for which the European Communities could face suspension of concessions. . . . (§ 6.12, p. 37)

The DSB identifies two principles—that nullification or impairment is measured by trade and only by trade between disputants.

## **Reciprocity**

The “rectangles” methodology also adheres to the WTO/GATT principle of reciprocity, though this is not a principle the DSB explicitly used to determine damages in *Hormones*. The reciprocity principle is that in negotiating tariff reduction, each country should make *equivalent* tariff concessions. Tariff concessions are not calculated based on estimates of how distorting each country’s tariff is but rather all tariffs are eliminated or harmonized by a specific amount. For example, all agricultural tariffs are reduced (or eliminated) by a specific percentage of the entire volume of trade. Portions of that trade are not exempt depending on measures of producer and consumer welfare. Therefore, since tariff concessions are negotiated as blocks of trade (i.e., rectangles), then it is logical and consistent with WTO principles that the suspension of tariff concessions should also be calculated as rectangles. Calculating trade damages as revenue losses may be a clearer demonstration of the reciprocity principle than calculating damages as welfare losses.

## **Framing and Estimating the Counterfactual**

In *Hormones*, the DSB panel estimated impairment as the value of EU imports from the United States and Canada foregone or “lost” because of the ban. The counterfactual is the volume that *would have been* imported had the EU hormone ban not been in place. Two types of beef were affected by the EU ban: High-Quality Beef—used by hotels and restaurants, and Edible Beef Offal—hearts, livers, kidneys, tongues, sweetbreads, and tripe, among other by-products. These are two distinct products trading in distinct markets and with different trade regimes.

## **High Quality Beef**

The HQB counterfactual was relatively simple to construct. Because the EU maintains a Tariff-Rate Quota (TRQ) on HQB imports from the United States and Canada and because the overquota tariff is prohibitive, the in-quota volume provides an upper bound. Impairment was measured as the difference between the in-quota TRQ volume and the actual volume of HQB imports during a representative



period multiplied by a representative HQB price. The DSB also needed to determine the U.S. and Canadian shares of the TRQ: It ruled 92% and 8% respectively.

### **Beef Offal**

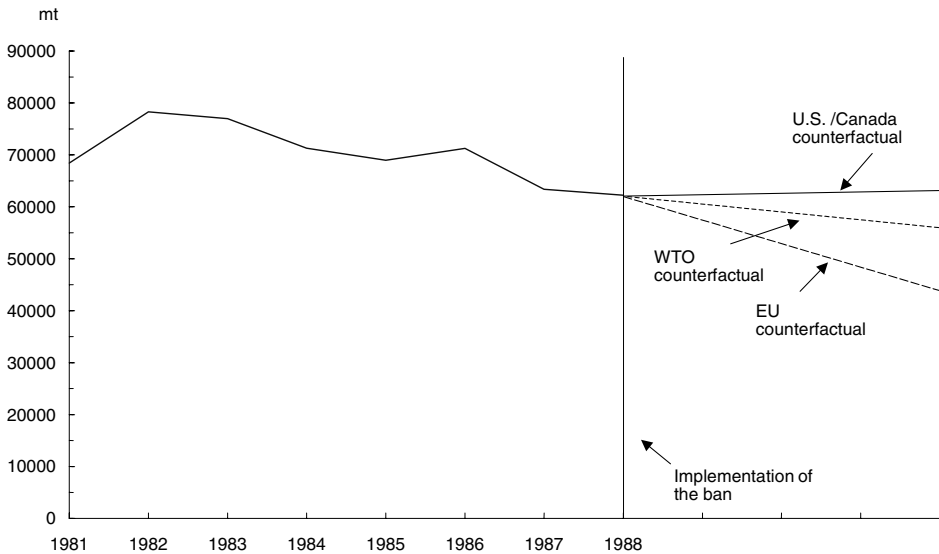
The DSB's counterfactual for beef offal was more difficult to construct. The DSB Panel's counterfactual was the value of offal that would have been imported had the ban been removed on May 13, 1999. The Panel's choice of date is important. Although the ban had been in dispute for over 11 years, it was only in 1998 that the DSB found it to be inconsistent with the EU's WTO obligations. The DSB ruled that the ban was only in violation of WTO obligations once the 15-month implementation period had passed, that is, on May 13, 1999. Only then could one argue that the ban improperly impaired trade.

The complainants assumed a different counterfactual. They presumed the EU ban was in violation retroactively. They asked what the level of imports would have been had the EU never imposed the ban. The two counterfactuals might be similar if the ban had no long-term impact on the structure of the EU beef offal market. Removing the ban would simply allow trade to return to its long-term trend. However, if after 11 years, structural changes were induced by the ban, then removing it would shift trade to a different long-run trend and distribution of trade.<sup>9</sup>

Prior to the ban, EU beef offal imports faced a 4% tariff and some routine health and safety standards: There was relatively free trade in beef offal. However, there was a declining trend in EU beef offal imports from the United States in some of the years before the ban took effect (figure 3). The complainants, the EU, and the DSB all provided different theories on the reasons and importance of this decline in their respective counterfactuals.

As shown in figure 3, the United States and Canada assumed that there would not be a significant decline in beef offal imports if the ban had not been implemented; the downward trend in imports in the time period before the ban was only a short-run phenomenon.<sup>10</sup> The EU assumed that there would be a significant decline in offal imports even if the ban had not been implemented due to non-ban-related factors. The WTO counterfactual turned out to be between the United States and EU counterfactuals, but the method the DSB panel actually used to estimate this counterfactual was somewhat complex. The panel first assumed that there was a drop in EU beef offal consumption after the ban was imposed (1989) that was unrelated to the ban itself. The DSB then estimated the counterfactual level of beef offal imports from the EU by using a simple linear trend extrapolating import data for the 8 years before the ban was enacted (1981–88) to the 3 years after the ban (1989–91). The difference between the trend import level and the actual imports of beef offal between 1989 and 1991 was assumed to be the lost imports caused by the ban. Adding the annual average of this difference to actual beef offal imports<sup>11</sup> between 1995 and 1997 would be the counterfactual level of imports. This counterfactual is equivalent to an 18.4% drop in EU beef offal consumption (from the 3-year period before the ban) that can be attributed to non-ban-related factors. The panel assumed that the volume of exports would have declined in proportion to this decline in consumption if not for the ban.

At first, this approach to determining the counterfactual level of imports seems plausible. Since there was a declining trend in EU beef offal imports in some

**Figure 3. EU beef offal imports from the United States**

of the years before the ban took effect, it would be reasonable to assume that such a trend would continue even if the ban were not enacted. Moreover, using a least-squares linear regression may not be the most sophisticated approach to forecasting, especially when estimating imports only a few years after the ban took effect, but such forecasts are the best linear unbiased estimates (assuming that classical assumptions hold.) One objection to the panel's method is that imports should be forecast as the residual of domestic production and consumption rather than as a reduced-form function of import unit values. This method is almost the opposite of the panel's methodology, which first forecast a decline in imports then translated this change into an equivalent change in consumption.

### The DSB's Determination of Facts

In *Hormones*, the DSB panel acted like arbitrators in other formal dispute settlement systems. The DSB used the same measure of economic damage: It used the disputants' final offers and the methodologies employed as a source of information about the facts, and determined damages by weighting the conflicting claims by their relative plausibility.

In the arbitration hearing, the DSB panel posed questions about production, consumption, and trade. One issue the panel focused on was whether there was a significant decline in EBO consumption unrelated to the ban. The panel sought evidence about the plausibility of the competing counterfactuals: It was not especially interested in the economic models used.

The DSB concluded that there was a *prima facie* case that there was some decline in beef offal consumption unrelated to the ban and gave relatively more weight to the EU counterfactual (European Communities, 1999a). On other questions, the DSB accorded more weight to the U.S. counterfactual. Based on data from the arbitrators' report, table 1 compares the damage estimates of the DSB, the United States, and the EU.

Table 1. Damage estimates from the DSB, the United States, and the EU

		WTO (DSB)	U.S.	EU	Average of U.S. and EU estimates
High Quality Beef (HQB)					
(1)	Volume of lost trade	11,500 mt	11,500		
(2)	Adjustment factor	92%	100%		
(3)	Price	\$5,342	\$5,342		
(4) = (1) × (2) × (3)	Value of lost trade	\$56,518,360	\$61,433,000	\$33,488,129	
(5)	Gross value of current imports	\$31,804,779		\$31,804,779	
(6)	Adjustment factor	75%		100%	
(7) = (5) × (6)	Adjusted value of current imports	\$23,853,584	—	\$31,804,779	
(8) = (4) – (7)	Net value of lost trade	\$32,664,776	\$61,433,000	\$1,683,350	\$32,399,850
	Difference between WTO and EU-U.S. average				–0.81%
Edible Beef Offal (EBO)					
(1)	Volume of lost trade	65,568 mt	70,989	45,253	
(2)	Adjustment factor	81.6%	100%	75%	
(3)	Price	\$1,689	\$1,689	\$1,684	
(4) = (1) × (2) × (3)	Value of lost trade	\$90,367,391	\$119,900,000	\$56,795,838	
(5)	from promotion	—	\$20,100,000	—	
(6)	Value of current imports	\$2,460,759	—	\$2,460,759	
(7)	Adjustment factor	75%		100%	
(8) = (6) × (7)	Adjusted value of current imports	\$1,845,569	—	\$2,460,759	
(9) = (4) + (5) – (8)	Gross value of lost trade	\$88,521,822	\$140,000,000	\$54,335,079	
(10)	Adjustment Factor (pet food)	95%		95%	
(11) = (9) × (10)	Net value of lost trade	\$84,095,731	\$140,000,000	\$51,618,325	\$85,809,163
	Difference between WTO and EU-U.S. average				2.04%
Total (EBO + HQB)					
		\$116,760,507	\$201,433,000	\$53,301,675	\$117,367,338
	Difference between WTO and EU-U.S. average				0.52%

For High-Quality Beef, the United States argued that damages should consist of 100% of the TRQ. The EU argued that the United States would not be able to fill the TRQ even if the ban were not in place and only used an approximate level of trade. The DSB assumed that the United States was essentially correct and awarded (92%) of the damages in this category that they asked for. Regarding deducting the value of current imports, the United States argued that the counterfactual level of damages should not be reduced by the level of current imports since the EU planned to restrict all imports of beef, even hormone-free beef, in the future. The EU argued that damages should be reduced by the full amount of current trade. The DSB found the EU argument more plausible but only reduced imports by 75% of current trade.

For Edible Beef Offal, the United States and the EU came up with very different estimates of lost trade. The United States claimed a larger value than the EU and assumed there would be no significant, long-term decline in the beef offal consumption in the EU if not for the ban. The EU used a much lower estimate of the beef offal trade and made a significant (25%) reduction in imports because of non-ban-related factors. As shown in figure 3, the DSB seems to have split the difference. They granted that the actual level of trade was closer to the U.S. estimate but that there was a non-ban-related decline in EBO consumption in the EU of about 14%. The EU also argued that 5% of the beef offal imports go into the production of pet food and since pet food is not prohibited by the ban, the United States could not claim compensation for such trade.

The DSB damage estimates are almost exactly equal to a simple average of the two parties' estimates as shown in the right hand column of table 1. This result could be chalked up to coincidence except that in other arbitration cases, the claim has been made that arbitrators "split the difference" when determining damages. We emphasize that it is naïve to assert that the DSB simply "split the difference" when calculating damages. However, let us assume that the DSB is indeed weighing each of the party's estimates by the relative strength of the arguments presented. Assuming an equal tendency for the plaintiff and defendant to overestimate and underestimate the damage amounts, respectively, it is reasonable that the final weight for the two parties' estimates would come close to 50–50 and the damage award would approach an average of the two estimates.

## Conclusion

When calculating damages, the DSB desires a simple, straightforward method. The arbitration panel is more interested in the arguments presented to defend damage estimates than in the choice of economic model or formula used. The panel avoids complex methodologies that can be easily subject to dispute or criticism, even if that means sacrificing some accuracy in determining the exact amount of damages. The DSB uses a methodology based on calculations of gross trade between the two countries that are parties to the dispute in question ("rectangles"). Welfare measurements ("triangles") or trade involving third-party countries are not considered. Using such measurements would only expand controversy and undermine the entire dispute settlement process. For this reason, the "rectangles not triangles" is the methodology adopted in most arbitration systems.

The DSB, like other formal dispute settlement systems, must strike a balance between awarding damages that are too high and encouraging litigation and awarding damages that are too low and encouraging violations of WTO agreements. In *Hormones*, a welfare analysis would result in much lower level of impairment suffered by the United States as a result of the EU hormone ban (because third party trade reduced the amount of “lost” exports) than an analysis of the gross trade between the United States and EU. However, if the WTO actually calculated damages using a welfare analysis or “triangles” methodology, this may act as an incentive for countries to violate their WTO agreements. The damage that a country may be forced to pay in such instances may be a small price compared with the economic and political costs of removing the policy.<sup>12</sup> There is also a problem with using a calculation of gross trade or “rectangles” methodology, since it may result in a flood of new litigation. If a country thought it might get the full value of the lost trade that existed before the WTO-illegal policy was put into effect, there would be much greater incentive to petition the WTO for damages and are less incentive to settle.

While it is true that our hypothesis does seem to correspond to certain facts of trade dispute cases (the *Hormones* case in particular) and fits theories of law and economics literature nicely, it would be rash to assume that our hypothesis is correct. It should be subject to more rigorous testing as more trade dispute cases go through the arbitration process for damages. However, if this hypothesis were true, this would have important implications for those calculating damages for trade dispute cases in front of the DSB. A party to a dispute should concentrate more on constructing good qualitative arguments to support its estimates and developing answers to questions from a DSB panel that would likely support these arguments. Less time should be given to developing complex methodologies or economic models to forecast trade, since the DSB will likely favor a more simple, straightforward approach in any case. In the least, we hope that this paper will encourage further research on trade dispute cases or at least get people to think about these cases a little differently.

## Endnotes

<sup>1</sup>The analysis follows the textbook, partial equilibrium analysis of a tariff in a small country, for example, Salvatore (90 ff.), Corden (5 ff.)

<sup>2</sup>The WTO Agreements are agreements between member countries. Citizens of member countries cannot initiate disputes against their own governments or against foreign member governments under the WTO.

<sup>3</sup>*Wickard v. Filburn* 317 U.S. 111 (1942), page 128.

<sup>4</sup>Formally, the impact of a single agent,  $\epsilon X$ , may be trivial, but, in aggregate, the impact of  $N$  agents,  $N\epsilon X$ , is substantial.

<sup>5</sup>*Wickard* was a radical departure from Court decisions of the mid-1930s. It established an important precedent; in particular, its expansive interpretation of the commerce clause was invoked to override segregation laws in the Southern states—*Heart of Atlanta Motel, Inc. v. United States*, 379 U.S. 241 (1964), *Katzbach v. McClung*, 379 U.S. 294 (1964).

<sup>6</sup>Filburn's fine of \$117.11 implies a wheat price elasticity of demand of  $-1.93$ . Wheat demand is generally regarded as inelastic: Schultz estimated it to be  $-0.08$ , but this was for wheat for human consumption. In *Wickard*, the USDA contended that it is the marginal demand for wheat as feed that matters. Including the cross-elasticities of substitute feeds yields a more elastic demand for wheat, even an elastic demand. So, the fine imposed may have approximated the damage.

<sup>7</sup>The value of the vertical rectangle is \$225.86, which equals the value of the horizontal rectangle when the elasticity equals  $-1$ .

<sup>8</sup>Bloom also found that there was a substantial amount of unexplained variance in the arbitrators' decisions.

<sup>9</sup>If the counterfactual employed by the panel in *Hormones* is a precedent for future disputes, the question of whether this will result in a different damage estimates than if the noncompliant measure never occurred is deserving of further study.

<sup>10</sup>Interestingly, the U.S. and Canada used different economic models to estimate the level of imports but came up with about the same number.

<sup>11</sup>The EU was importing some hormone-free beef offal after the ban was implemented and claimed that some hormone-treated beef offal was being imported for use in pet food.

<sup>12</sup>Of course, it is not the "country" that pays. It is the exporter who must pay the higher tariff and the consumer who may have to pay a higher price. Indeed, most economists would say that the suspension of tariff concessions as a way of penalizing countries is not only inefficient but results in a net welfare loss.

## References

- Ashenfelter, O. "Evidence on U.S. Experience with Dispute Resolution Systems." Working paper, Industrial Relations Section No. 185, Princeton University, 1985.
- Ashenfelter, O., and D.E. Bloom. "Models of Arbitrator Behavior: Theory and Evidence." *Amer. Econ. Rev.* 74(March 1984):111–24.
- Bazerman, M., and H. Farber. "Arbitrator Decision Making: When are Final Offers Important?" *Indust. Labor Rel. Rev.* 39(Oct. 1985):76–89.
- Benson, B.L. "7500 Arbitration: Literature Review." *Encyclopedia of Law and Economics*, (<http://encyclo.fndlaw.com/lit/7500art.html>).
- Bloom, D. "Empirical Models of Arbitrator Behavior Under Conventional Arbitration." *Rev. of Econ. and Stat.* 68(Nov. 1986):578–85.
- Cooter R., and D. Rubinfeld. "Economic Analysis of Legal Disputes and Their Resolution." *J. of Econ. Lit.* 27(Sept. 1989):1067–97.
- Corden, W.M., *The Theory of Protection*. Oxford: Oxford University Press, 1977.
- European Communities–Measures Concerning Meat and Meat Products (Hormones): Decision by the Arbitrators, World Trade Organization, WT/DS26/ARB, July 1999a, ([http://www.wto.org/english/tratop\\_e/dispu\\_e/26arb.pdf](http://www.wto.org/english/tratop_e/dispu_e/26arb.pdf)).
- European Communities–Regime for the Importation, Sale, and Distribution of Bananas: Decision by the Arbitrators, World Trade Organization, WT/DS27/ARB, April 1999a ([http://www.wto.org/english/tratop\\_e/dispu\\_e/1735d.pdf](http://www.wto.org/english/tratop_e/dispu_e/1735d.pdf)).
- Farber, H. "Splitting the Difference in Interest Arbitration." *Indust. Labor Rel. Rev.* 35(Oct. 1981):70–77.
- Gunderson, M., "Economic Aspects of Interest Arbitration." Ontario Economic Council Paper Series, 1983.
- Posner, R., *Economic Analysis of Law*, 2nd ed. New York: Little Brown and Company, 1977.
- Salvatore, D., *International Economics*, 2nd ed. New York: McGraw-Hill, 1984.
- Schultz, H., *The Theory of Measurement and Demand*. Chicago: The University of Chicago Press, 1938.
- Stewart, T., ed., *Agricultural Sanitary and Phytosanitary and Standards Report: The WTO Beef Hormones Decisions*, November 1997.
- Wickard v. Filburn*, 317 U.S. 111 (1942).